IOWA STATE UNIVERSITY
Digital Repository

Volume 44 | Number 2 Article 8

10-2016

Experiences with Campus Digital Photographs @ MSU Archives

Ed Busch
Michigan State University, buschedw@msu.edu

Mixed Media: Working with Audio and Visual Materials

Assistant Editor: Heather Fox, University of Louisville. Contact Heather at heather.fox@louisville.edu if you would like to guest author a column or have a good idea to share.

Experiences with Campus Digital Photographs @ MSU Archives

By Ed Busch, Michigan State University

Since coming to Michigan State University (MSU) Archives and working with born-digital photographs, I have pursued a number of workflows and tools to improve processing and to provide access. This short article will present some of the experiments and trials I have conducted with these digital photographs as well as where I am heading with these trials now.

This work first became a focus for me when I processed a collection of university records from one of the MSU colleges during 2008–2009 that included over 400 magnetic and optical media items, including floppies, CD/DVDs, and Zip and Jaz disks. Floppies can hold from 100K to 1.5 MB of data, while CD/DVDs can hold from 650MB to 4.7 GB. Iomega Zip and Jaz disks were popular storage options in the mid- to late 1990s. Zip disks initially stored 100MB but later handled up to 750MB; Jaz disks initially held 1GB, which later increased to 2GB.

At that time, the MSU Archives had neither a policy nor a workflow in place for handling these electronic materials, nor all the hardware or software needed to access the various storage types and file formats. We had already set up 3½" and 5¼" floppy drives. We acquired Zip and Jaz drives from eBay. Many of the CDs were Kodak Photo CDs, created when rolls of film were processed. Images were then scanned by the Kodak film processor at a resolution of 2200dpi and burned to the CD in the PCD file format. Reading these files accurately and completely requires special software. We purchased an application called pcdMagic to read them and convert them to Tagged Image File Format (TIFF), an uncompressed file format frequently selected as the image format for preservation systems.

While processing this media, I created a workflow for their handling that used a spreadsheet as a log for tracking details about the media. I assigned an identifier to each piece of media and then recorded this identifier along with media type, description (including date), physical location, file location, and accession number in the spreadsheet.

(Continued on page 28)

	A	B	C	D	E	F	G	H	I
201	CHM-00199	Floppy 3.5	Publications Office (2001)		R:\UA 15 13 CHM\UA06247	UA 6247			
202	CHM-00200	Floppy 3.5	Unlabeled (1998)		R:\UA 15 13 CHM\UA06247	UA 6247			
203	CHM-00201	CD	MSU_LLC Event		R:\UA 15 13 CHM\UA06416	UA 6416			
204	CHM-00202	PHOTO CD	PCD0886 (PCD format)		R:\UA 15 13 CHM\UA06416	UA 6416			
205	CHM-00203	CD	Conley2.10.12.07 (CHM Alumni Reception at U. Club)		R:\UA 15 13 CHM\UA06416	UA 6416			
206	CHM-00204	CD	CHM Alumni Relations - Alumni Board Group Photo 10-12-2007		R:\UA 15 13 CHM\UA06416	UA 6416			
207	CHM-00205	CD	MSU CHM Coach P Golf Outing Brochure 6-24-2001		R:\UA 15 13 CHM\UA06247	UA 6247			
208	CHM-00206	CD	White Coat Ceremony (2004?)		R:\UA 15 13 CHM\UA06247	UA 6247	Disk errors in TIF folder		
209	CHM-00207	CD	White Coat Ceremony (5-10-2004)		R:\UA 15 13 CHM\UA06247	UA 6247	copy of CHM-00206. Has disk errors in TIF folder too.		
210	CHM-00208	CD	CHM Commencement 5/15/04		R:\UA 15 13 CHM\UA06247	UA 6247	Disk errors in TIF folder		
211	CHM-00209	CD	The Follies 4/23/04		R:\UA 15 13 CHM\UA06247	UA 6247			
212	CHM-00210	CD	CHM Commencement 2003		R:\UA 15 13 CHM\UA06247	UA 6247			
213	CHM-00211	PHOTO CD	5/21/03		R:\UA 15 13 CHM\UA06247	UA 6247			
214	CHM-00212	PHOTO CD	5/21/03		R:\UA 15 13 CHM\UA06247	UA 6247			
215	CHM-00213	PHOTO CD	6/3/2003		R:\UA 15 13 CHM\UA06247	UA 6247		Click to adc	
216	CHM-00214	PHOTO CD	Follies 2002		R:\UA 15 13 CHM\UA06247	UA 6247	This is really CHM221, moved		
217	CHM-00215	PHOTO CD	CHM Commencement 2001 #1-50 3 of 3 / MMSA Banquet 2001 #51-86 2 of 2		R:\UA 15 13 CHM\UA06247	UA 6247			
218	CHM-00216	PHOTO CD	CHM Commencement 2001 (#1-42, 79-100) 2 of 3 / MMSA Banquet 2001 (#43-78) 1 of 2		R:\UA 15 13 CHM\UA06247	UA 6247			

This spreadsheet tracked mixed media types from the College of Human Medicine accession. Image courtesy of Ed Busch.

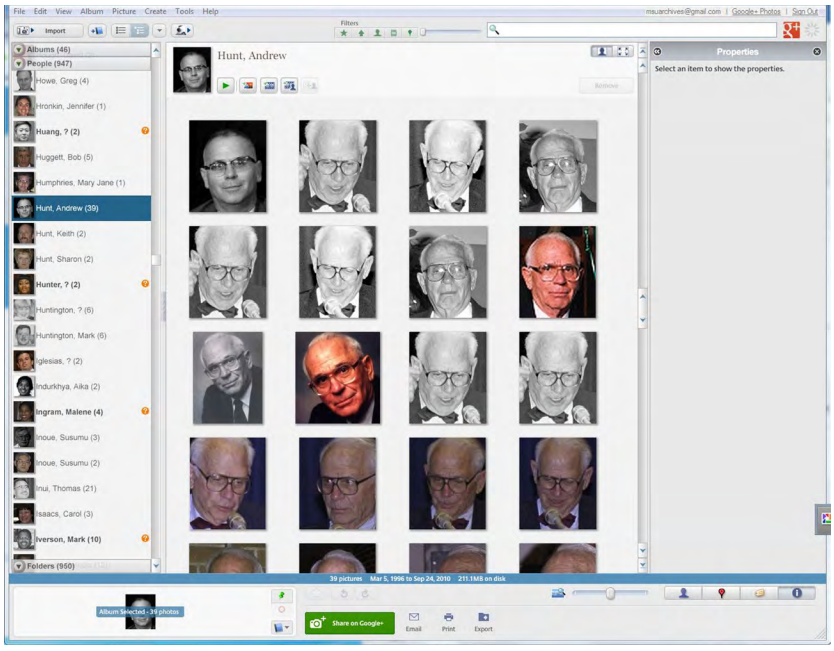
Digital Photographs @ MSU

(Continued from page 27)

In 2010, the files were virus checked, transferred off of the media, and stored to a networked storage device (over 42,000 files, 220 GB). At this time, it became important to me to find a way to search images to handle requests from, and ideally share images with, creating departments and/or to provide an external access method. I was familiar with iPhoto on the Mac and needed something to use on our office PCs that had common image organizer features such as thumbnail previews, album creation, tags, and an access method (export, e-mail, printing, etc.). I was also looking for something free or inexpensive that was easy to install and use (for archivists and the public) to provide access control. A big plus was facial recognition and metadata handling functions.

Picasa was one of the Google apps I tried. It was free and came with 1 GB of online storage for public access. I selected Picasa primarily for its facial recognition feature and its support of Adobe XMP and EXIF metada0ta. I was very interested in facial recognition and its potential power to help identify people in photographs more efficiently.

For my implementation at MSU, I used Picasa's folder manager setting to specify folders to scan for updates (new images) and also turned on face detection. Picasa's face detection works by identifying faces in images and then trying to mark other images with similar characteristics.



This snapshot from Google's Picasa shows images identified by facial recognition of former dean Andrew Hunt. Image courtesy of Ed Busch.

Face detection identified candidates, but took many decisions by me to sort the good from bad matches. Fortunately, many of the photo files included names in their titles, which helped identify faces. I found Picasa's facial recognition very useful for quickly finding known people in many photographs. I continued to use Picasa on my desktop to search and access files, but, due to online limitations, I never implemented the public search. Updates by Google for the Picasa application became less frequent.

During this time, we conducted ongoing discussions with our campus photographers to come up with a workflow for transferring and managing their photographs. The campus photographers unit transferred DVDs with over 39,000 photographs (665 GB) to the archives. They also had many more photographs on local and networked storage to transfer at a future date. I learned that they were using Adobe Lightroom on their desktops for editing and tagging. Completed photos were uploaded to their Zenfolio website for public access and purchase. I experimented with Lightroom in our environment, but as it lacked facial recognition, I held off on implementing it into our workflow.

In 2015, the campus photography unit suggested that we begin using Zenfolio with the thought that they could transfer their account holdings to our account in the future and that we could then download and ingest their content into our preservation environment. They were just beginning to look for a new product, as Zenfolio lacked the robust search functionality that they needed. We signed up for a Zenfolio account and uploaded the contents of their DVDs in our holdings to our Zenfolio account (663 GB). During that year, Adobe also released the first version of Lightroom to include face recognition. After many Google searches and experiments with various tools, I was able to transfer my Picasa data with its facial recognition tags into Lightroom. At least, the metadata transferred in a fashion.

Using Lightroom, I worked on Picasa metadata cleanup, experimented with adding metadata (keywords and/or captions), and then uploaded to Zenfolio. Some of the

keywords were added using Lightroom's facial recognition. I continued the Zenfolio upload to provide a basic search function for our staff since an access portal for our MSU Trusted Digital Repository (TDR) does not exist yet.

For access to the MSU Archives digitized images (scanned images), we have been using a Flickr account and an MSU-developed platform, KORA. Flickr is an image- and video-hosting website. KORA is an open source, database-driven, online digital repository application for complex multimedia objects (text, images, audio, video) created by MATRIX at MSU (kora.matrix.msu.edu). We do not currently put our born-digital images on these sites, and our Zenfolio account is currently set to private for internal use only. In the near future, we plan on implementing Zenfolio's storefront since many of our images are popular with the general public and can generate a nice revenue stream for the archives.

So, where does this bring me to now? I am still experimenting and still need to develop a good workflow for born-digital photographs coming in from campus units and particularly from university photographers. The

(Continued on page 30)

Hunter Doherty "Patch" Adams attended a College of Human Medicine event in 1998. Image courtesy of Michigan State University Archives and Historical Collections.

Digital Photographs @ MSU

(Continued from page 29)

workflow needs to include processing photographs (weeding), adding metadata (facial recognition), and preserving the images in our TDR. Of particular importance is getting a good plan worked out with the university photographers on what they should send to the archives. We do not want everything, but what we want is not necessarily what the photographers think is important. The photographers usually focus on the people at an event, while the archivist may also be interested in details around the people, such as the building interior or exterior.

We also need to monitor where the university photographers are going with their digital asset management tool search and selection. Zenfolio provides a basic search function and does provide a storefront when we are ready to go in that direction. It does not provide great organizing options nor is data entered in Zenfolio available when downloaded (metadata and captions). The MSU photographers are currently looking at digital management tools such as SmugMug and others.

Implementing a solid workflow for the handoff of born-digital photographs between the campus photographers and the university archives and providing a public access portal will have a positive impact on our acquisitions. It will also be a great boon for our outreach program at the MSU Archives. Adding useful keywords by working with the photographers and using techniques such as facial recognition will make access that much easier for researchers, the general public, and the archives staff. The MSU Archives' digital images, both born digital and digitized, are used for many different types of publications (paper and digital print, television, public events and lectures, etc.), providing a great opportunity for showcasing our unique digital images to the public. I hope to share more at local and regional conferences as we work through our goals for preserving and sharing our university photographs.

MSU's "Sparty" mascot with Dean Marsha Rappley at the College of Human Medicine's 40th Anniversary event in 2004. Image courtesy of Michigan State University Archives and Historical Collections.